

Physical Fitness and Body Composition in Transgender vs. Cisgender Individuals: A Systematic Review and Meta-Analysis

Sofia Mendes Sieczkowska^{1,2,3}, Bruna Caruso Mazzolani^{1,2}, Danilo Reis Coimbra³, Igor Longobardi^{1,2}, Andresa Rossilho Casale^{1,2}, José Davi F.V.M.P. da Hora⁴, Hamilton Roschel^{1,2}, Bruno Gualano^{1,2}

¹Applied Physiology and Nutrition Research Group –School of Physical Education and Sport and Faculdade de Medicina FMUSP, Universidade de Sao Paulo, Sao Paulo, SP, Brazil

²Center of Lifestyle Medicine; Laboratory of Assessment and Conditioning in Rheumatology; Hospital das Clínicas HCFMUSP, Faculdade de Medicina FMUSP, Universidade de Sao Paulo, Av. Dr. Arnaldo, 455, 3º andar, São Paulo, SP 01246- 903, Brazil

³Federal University of Juiz de Fora, School of Physical Education and Sports, Department of Sports, Juiz de Fora, MG, Brazil

⁴School of Medicine, University of Sao Paulo, Institute of Psychiatry, Clinical Hospital, São Paulo, SP, Brazil.

Corresponding author: Bruno Gualano

Address for correspondence: Center of Lifestyle Medicine, School of Medicine, University of Sao Paulo, Sao Paulo, SP, Av. Dr. Arnaldo, 455, 3 andar, 01246-903 São Paulo, Brazil; E-mail: gualano@usp.br

ORCID iDs

Sofia Mendes Sieczkowska - 0000-0002-4869-318X

Bruna Caruso Mazzolani - 0000-0002-4687-3128

Danilo Reis Coimbra - 0000-0002-6055-0667

Igor Longobardi - 0000-0001-6669-9553

Andresa Rossilho Casale - 0000-0002-4281-3427

José Davi F.V.M.P. da Hora - 0009-0002-4267-8402

Hamilton Roschel - 0000-0002-9513-6132

Bruno Gualano - 0000-0001-7100-8681

Funding

Authors were supported by São Paulo Research Foundation (FAPESP) and Coordination for the Improvement of Higher Education Personnel (CAPES).

Competing interests

All authors confirm that there are no conflicts of interest associated with this work.

Contributorship

SMS and DCR searched studies in the databases, ran statistical analysis, elaborated results and drafted the manuscript. SMS, BCM, DCR, IL, ARC and JFVMPH extracted data. BCM and IL evaluated risk of bias, edited and revised the manuscript. BG conceived the study design and drafted the manuscript. HR conceived the study design and revised the manuscript. All authors contributed to critical revision of the report for important intellectual content.

Acknowledgements

None to declare.

Data sharing statement

All data produced in the present study are available upon reasonable request to the authors.

Ethics approval

Not applicable.

Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Abstract

Objective: To compare fitness and body composition between transgender and cisgender individuals.

Design: Systematic review and meta-analysis.

Data sources: PubMed, Web of Science, Embase, and SportDiscus databases were searched in June 2024, supplemented by manual citation reviews.

Eligibility criteria: Inclusion criteria comprised studies of transgender individuals comparing physical fitness/body composition pre-to-post gender-affirming hormone therapy (GAHT) or vs. cisgender controls, with quantitative outcomes reported.

Results: Fifty-one studies (6,434 participants) were analyzed. Transgender women (TW) exhibited comparable fat mass to cisgender women (CW), but higher than cisgender men (CM). TW showed greater lean mass than CW, but lower than CM. Upper- and lower-body strength were similar between TW and CW, but lower than CM. TW and CW had similar VO₂ peak, but TW exhibited lower values than CM. Transgender men (TM) exhibited similar fat mass to CW, but higher than CM. TM showed higher lean mass than CW, but lower than CM. Upper-body strength was higher in TM than CW, but lower than CM. GAHT in TW increased fat mass, reduced lean mass and upper-body strength, with no differences in lower-body strength over 1–3 years. TM demonstrated reduced fat mass and increased lean mass, upper- and lower-body strength post-GAHT. Risk of bias was moderate for most studies, with limited observations for specific outcomes (e.g., VO₂ peak).

Conclusion: While TW exhibited higher lean mass than CW, their physical fitness was comparable. Current evidence is limited but does not justify blanket bans based on assumptions of inherent athletic advantages for TW over CW.

Key-words: gender-affirming hormone therapy; performance; women; transsexual

WHAT IS ALREADY KNOWN ON THIS TOPIC

- The inclusion of transgender women in female sports categories remains a highly contentious subject.
- Existing studies suggest that gender-affirming hormone therapy (GAHT) alters body composition in transgender individuals, but evidence on functional performance outcomes (e.g., strength, endurance) remains inconsistent.
- Policies advocating blanket bans on transgender women (TW) in female sports often cite residual advantages from prior testosterone exposure, despite limited empirical support for sustained performance disparities post-GAHT.

WHAT THIS STUDY ADDS?

- This systematic review and meta-analysis synthesizes data from 51 studies (6,434 participants), demonstrating that, while TW showed higher absolute lean mass than cisgender women (CW), there are no significant differences in upper/lower- body strength or VO₂ peak after 1–3 years of GAHT.
- These findings challenge the validity of blanket bans predicated on assumptions of inherent athletic superiority of TW over CW.
- Transgender men (TM) exhibit body composition and strength metrics intermediate between CW and cisgender men (CM) post-GAHT.
- Critical research gaps are identified, including a lack of long-term GAHT data, underrepresentation of transgender athletes, and inconsistent controls for confounders (e.g., training history, puberty blocking). Future studies must prioritize sport-specific performance metrics and longitudinal designs to inform equitable policies.

Introduction

The question of whether transgender women (TW) should be permitted to compete in female sports, and under what conditions, remains a subject of intense debate. The rationale for sex-segregated competition is rooted in ensuring equitable opportunities for cisgender women (CW), prompting proposals that transgender (or intersex) athletes should be included only if their participation does not disproportionately disrupt competitive fairness [1, 2, 3].

However, empirical evidence challenges initial concerns that TW would dominate women's sports, largely due to the physiological effects of testosterone suppression therapy [1]. In fact, TW remain underrepresented in elite athletics. For example, Laurel Hubbard, the first openly transgender woman to compete in the Tokyo Olympic Games (2021), participated in weightlifting but did not advance to medal contention, underlining the lack of dominance by transgender athletes in practice.

The International Olympic Committee (IOC) recently established a framework prioritizing fairness, inclusion, and nondiscrimination for athletes with diverse gender identities and sex variations [2]. This approach rejects blanket bans based on gender identity, advocating instead for sport-specific eligibility criteria informed by evidence. Critics, however, argue that the framework relies on insufficiently developed research and impractical case-by-case assessments, potentially compromising protections for cisgender female athletes. Lundberg et al. (2024) [3] contend that the IOC's "no presumption of advantage" principle overlooks studies suggesting that transgender women retain muscle mass, strength, and other physical advantages over cisgender women even after testosterone suppression. Their argument hinges on the well-established physiological disparities between cisgender males and females, which confer inherent athletic advantages to males. Nonetheless, systematic reviews comparing transgender women (post-hormone therapy) and cisgender women report inconsistent findings regarding performance and physical differences, highlighting the need for further research [1, 4].

To shed light on this topic, this systematic review with meta-analysis evaluates the current literature on the physical fitness and body composition of transgender individuals (i.e., TW and transgender men [TM]) relative to cisgender ones (i.e., CW

and cisgender men [CM]). By synthesizing available data, we aim to guide decision-makers and stakeholders in developing equitable, scientifically grounded policies for transgender participation in sports.

Methods

This systematic review was registered in PROSPERO (CRD42024562210) and follow the guidelines of the PRISMA 2020 statement [5].

Search Strategy

For comprehensive coverage of worldwide scientific production, the search was conducted in the following electronic databases: PubMed, Web of Science, Embase, and SportDiscus during the first week of June 2024. The search strategy used descriptors related to the population and outcomes as follows: Population (Transgender OR Transsexual OR Transgender Person OR Transsexualism OR Transgenderism OR “Transgender people” OR “gender reassignment procedure” OR “gender reassignment surgery” OR “gender change procedure” OR “Gender-affirming Treatment” OR transwoman OR transmen OR “gender affirming hormone therapy (GAHT)” OR transsexual OR “cross-sex hormone therapy” OR “trans people”) and Outcomes (physical fitness and body composition terms, see table S1 supplementary material). To select search descriptors, MeSH terms (Medical Subject Headings) were used. Additionally, a manual search of the references in selected studies were conducted to identify studies for inclusion. The search process was carried out independently by two researchers. In case of disagreement, a third reviewer was consulted.

Eligibility for Study Selection

All articles identified in the search were screened by two independent members of the research team using a 3-stage strategy: 1) Title and abstract screening, and 2) Full text review and 3) Conflict resolution / consensus phase. Any discrepancies were resolved through discussion, or third-party mediation, if required. The selection was made using the software Rayyan QCRI [6].

Study eligibility was based on the PECO criteria, described below:

Population: Individuals identifying as transgender.

Exposure: Gender-affirming hormone therapy.

Comparison: Before and after gender-affirming hormone therapy, and with cisgender individuals.

Outcomes: Measures of physical fitness and body composition.

No language restrictions were applied, and there were no restrictions on publication dates.

Data Extraction Process

A data extraction table was created and completed to gather key information from the selected studies, including author, study population, sample size, type of hormonal therapy, outcomes, and main findings.

Risk of bias

The risk of bias analysis was conducted using the 20-item AXIS (Appraisal tool for cross-sectional studies) [7], ROBINS-I (Risk Of Bias In Non-randomized Studies - of Interventions) [8] or ROB2 (Risk Of Bias 2) [9], depending on study design.

AXIS was used to assess the risk of bias in cross-sectional studies. It consists of 20 items, each rated as “Yes,” “No,” or “Don't Know.” Items 7 (“Were measures undertaken to address and categorize non-responders?”) and 14 (“If appropriate, was information about non-responders described?”) were marked as “Inapplicable” when the response rate was reported as 100%. As AXIS lacks a standardized scoring system, we adopted a method used in previous studies [10-12], assigning scores of 0 or 1 to each item to calculate an overall quality score. Specifically, for items 13 (“Does the response rate raise concerns about non-response bias?”) and 19 (“Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?”), a response of “No” was scored as 1, and “Yes” or “Don't Know” as 0. For items 7 and 14, “Yes” or “Inapplicable” were scored as 1, and “No” or “Don't Know” as 0. For all other items, “Yes” was scored as 1, and “No” or “Don't Know” as 0. Total scores were calculated by summing the individual item scores. Following recommendations [13], studies were classified as high quality (scores of 14–20; 70–100%), fair quality (scores of 12–13; 60–69.9%), or low quality (scores of 0–11; 0–59.9%).

ROBINS-I was used to assess the risk of bias in cohorts and quasi-experimental studies, and includes seven domains of bias: bias due to confounding, bias in participants' selection, bias in interventions classification, bias due to deviations from intended interventions, bias due to missing data, bias in outcomes measurement, and bias in reported result selection. Each domain is assessed through signaling questions to provide a comprehensive analysis of potential biases within the study. Each domain is assessed through signaling questions to provide a comprehensive analysis of potential biases within the study.

ROB2 was used to assess the risk of bias in randomized controlled trials (RCTs) and consists of five domains: bias arising from the randomization process, bias due to deviations from intended interventions, bias due to missing outcome data, bias outcome measurement, and bias in reported result selection. Similar to ROBINS-I, it also incorporates a set of signaling questions to determine whether each domain poses a high, low, or unclear risk of bias.

Two independent members conducted this analysis. Any discrepancies were resolved through discussion or third-party mediation, if required.

Statistical analysis

The metacont function in RStudio was used to perform a comprehensive statistical analysis of the effects of hormone therapy on transgender individuals. This analysis involved aggregating data from studies comparing baseline and post-hormone therapy measures, focusing on the impact of hormone therapy on key physical fitness and body composition parameters. A random-effects model was employed to account for variability both within and between studies, enabling a robust estimation of the overall effect of hormone therapy on physical fitness outcomes in transgender individuals. The analyses were conducted only when there were at least three studies for each subgroup.

In addition, the metacont function was also used to compare the effects of gender-affirming hormone therapy between transgender and cisgender individuals. This analysis synthesized data from multiple studies examining the same physical fitness and body composition variables, aiming to quantitatively assess the magnitude and direction of treatment effects in transgender in comparison to cisgender individuals. This approach provided valuable insights into the impact of gender-affirming hormone therapy on physical fitness and body composition parameters. The analyses were

conducted only when there were at least three studies for each subgroup. The entire analysis was conducted using RStudio (version 4.4.0) with the meta statistical package [14].

Results

The database search identified 1,705 studies. Two studies published later were included after checking the citations of the included articles and other sources (Google Scholar, social media). After removing duplicates ($n = 638$), 1,067 publications were screened for inclusion. Of these, 595 were excluded based on title review and 284 after abstract review. The remaining 188 papers were selected for full-text reading, from which three were excluded due to the population, 48 for outcome, and 86 for “wrong publication type”. Therefore, 51 studies were included in the review [15-65] and 42 in the meta-analysis (Figure S1).

Methodological Characteristics of the Studies

The analysis included a total of 6,434 individuals: 2,926 TW, 2,309 TM, 551 CW, and 648 CM. Participants' mean age ranged from 14 to 41 years. Of the 51 studies reviewed, 44 focused on adults, while seven involved adolescents. Regarding study design, 22 were prospective cohorts, nine retrospective cohorts, 16 cross-sectional studies, three randomized controlled trials, and one quasi-experimental studies. Only 15 studies incorporated any form of physical activity (PA) assessment. Among these, two exclusively recruited amateur athletes, and another reported including only sedentary individuals.

Methods for assessing PA levels varied significantly: three studies used weekly METs (metabolic equivalents); three applied the Baecke questionnaire (mean scores); two reported the percentage or number of "active" participants; two documented weekly exercise frequency; two measured weekly exercise duration; one used the Sport Index (mean scores); one simply noted "regular physical activity" without further detail, and one reported including only sedentary individuals ($PAL < 2.0$).

Reviewed studies provided details on various hormonal therapies and related findings (Table 2). Among reported treatments, GAHT was used in 49 studies, either alone or in

combination with: Antiandrogens (18 studies, exclusively for TW), GnRH agonists (GnRHa) (six studies). GnRHa alone was used in two studies.

The cross-sex hormone therapies included estrogen formulations (estradiol valerate, 17 β -estradiol, estradiol gel, and transdermal estradiol); testosterone formulations (testosterone enanthate, testosterone undecanoate [administered intramuscular, at low or high dose]); antiandrogens (cyproterone acetate, spironolactone, and finasteride); GnRH agonists (triptorelin, leuprolide acetate or goserelin acetate).

Therapy duration varied widely, ranging from 3 months to 14 years, with most studies reporting following participants for 1 to 3 years of therapy. Also, of the 51 studies included, 11 included participants who had undergone gender-affirming surgery and only six reported the use of puberty suppression.

In 15 cross-sectional studies, information regarding naïve status to hormone therapy, adverse effects, and dropouts were deemed less relevant due to the study design. Therefore, of the remaining 36 studies included, 23 exclusively enrolled hormone-naïve individuals (one focused solely on naïve TW), five included non-naïve participants and eight did not report this data. Regarding adverse effects, 27 studies did not report, eight observed none and one study reported several adverse events in both TM (muscle/joint pain, mild hypertension, reduced fasting insulin, androgenic alopecia) and TW (depression, elevated prolactin, galactorrhea, transient liver enzyme elevations, hypertension, increased fasting insulin, and skin irritation). Moreover, 20 studies lacked dropout data, 11 reported none and four had significant attrition (≥ 20 participants) and one had minimal dropouts ($n=2$).

Transgender compared with cisgender individuals

Fat Mass

Of the 22 studies reporting this comparison, 20 were included in the meta-analysis. Two studies were excluded because they either lacked extractable data (due to graphical format) or did not report mean/median values (or other measures of central tendency) for the outcome of interest.

In the comparison between TW and CW, the average duration of cross-sex hormone therapy for TW was 2.82 ± 2.09 (range: 1-7) years. No significant differences were observed between groups [SMD: -0.22, 95% CI (-0.62; 0.19)] (Figure 1).

Average duration of cross-sex hormone therapy for TW was 3.25 ± 2.49 (range: 1-8) years when comparing with CM, with TW presenting higher fat mass [SMD: 0.85, 95% CI (0.50; 1.19)] (Figure 1).

Average duration of cross-sex hormone therapy for TM was 3.13 ± 3.09 (range: 1-10) years in the comparison with CW. No significant differences were observed between groups [SMD: -0.24, 95% CI (-0.64; 0.15)] (Figure 1).

Dataset used to compare TM and CM showed an average duration of cross-sex hormone therapy of 4.0 ± 4.62 (range: 1-14) years for TM. TM had a higher fat mass than CM [SMD: 0.90, 95% CI (0.50; 1.31)] (Figure 1).

Lean Mass

Of the 21 studies reporting this comparison, 18 were included in the meta-analysis. Three studies were excluded due to lack of extractable data (due to graphical format) or not reporting mean/median values (or other measures of central tendency) for the outcome of interest.

Cross-sex hormone therapy for TW was 2.80 ± 2.20 (range: 1-7) years on average when comparing with CW, and TW had a higher lean mass [SMD: 0.98, 95% CI (0.03; 1.93)] (Figure 2).

In the comparison between TW and CM, the average duration of cross-sex hormone therapy was 3.67 ± 2.84 (range: 1-8) years and TW exhibited a lower lean mass [SMD: -1.17, 95% CI (-1.58; -0.76)] (Figure 2).

Average duration of cross-sex hormone therapy for TM was 3.89 ± 3.69 (range: 1-10) years in the comparison with CW and TM exhibited a higher lean mass [SMD: 0.98, 95% CI (0.02; 1.94)] (Figure 2).

TM cross-sex hormone therapy duration average was 4.0 ± 4.62 (range: 1-14) years when comparing with CM, with TM exhibiting lower lean mass [SMD: -1.57, 95% CI (-2.57; -0.56)] (Figure 2).

Upper-body strength

All 12 studies reporting this outcome were included in the meta-analysis.

TW average duration of cross-sex hormone therapy was 3.17 ± 2.64 (range: 1-7) years. No significant differences were observed between groups when compared with CW [SMD: 0.41, 95% CI (-0.10; 0.92)] (Figure 3).

When comparing TW with CM, average duration of cross-sex hormone therapy was 3.86 ± 3.02 (range: 1-8) years. TW had a lower upper-body strength [SMD: -1.79, 95% CI (-2.55; -1.04)] (Figure 3).

In the comparison between TM and CW, the average duration of cross-sex hormone therapy for TM was 5.0 ± 4.58 (range: 1-10) years. TM had a higher upper-body strength than CW [SMD: 0.76, 95% CI (0.10; 1.42)] (Figure 3).

Average duration of cross-sex hormone therapy for TM was 9.0 ± 7.07 (range: 1-14) years when comparing with CM, and TM had a lower upper-body strength [SMD: -1.46, 95% CI (-1.94; -0.99)] (Figure 3).

Lower-body strength

Four studies reported this outcome and were included in the meta-analysis.

In the comparison between TW and CW, the average duration of cross-sex hormone therapy was 2.75 ± 2.87 (range: 1-7) years. No significant differences were observed between groups [SMD: 0.05, 95% CI (-0.74; 0.83)] (Figure 4).

Average duration of cross-sex hormone therapy for TW was 3.80 ± 3.42 (range: 1-8) years when comparing with CM, with TW exhibiting lower lower-body strength [SMD: -1.99, 95% CI (-2.66; -1.31)] (Figure 4).

VO₂ peak

All three studies reporting this outcome were included in the meta-analysis. For VO₂ peak, comparisons were only possible with TW, and average duration of cross-sex hormone therapy was 2.75 ± 2.87 (range: 1-7) years. No significant differences were observed between TW and CW [SMD: 0.07, 95% CI (-0.67; 0.82)] (Figure 5). TW exhibited lower VO₂ peak when compared to CM [SMD: -1.61, 95% CI (-2.69; -0.54)] (Figure 5).

Effects of cross-sex hormone therapy in transgender individuals

Fat mass

Twenty-three out of the 28 studies reporting this outcome were included in the meta-analysis. Five studies were excluded due to lack of extractable data (due to a graphical format) or lack of report on the duration of cross-sex hormone therapy.

First, we conducted the analysis regardless of whether the individuals were naïve or had undergone puberty suppression.

In TW, an increase in body fat mass was observed after one year of cross-sex hormone therapy compared to baseline [SMD: 0.44, 95% CI (0.27; 0.61)] (Figure S2). A similar effect was observed in studies with follow-up between one and two years [SMD: 0.59, 95% CI (0.22; 0.96)] and in those with a three-year follow-up [SMD: 0.70, 95% CI (0.19; 1.21)] (Figure S2).

When considering studies with hormone-naïve TW, the increase in body fat mass persisted up to one year of therapy compared to baseline [SMD: 0.50, 95% CI (0.36; 0.65)], between one and two years [SMD: 0.89, 95% CI (0.59; 1.18)], and after three years [SMD: 1.29, 95% CI (0.86; 1.73)] (Figure S3).

In TM, no increase in body fat mass was observed after one year of cross-sex hormone therapy compared to baseline [SMD: -0.15, 95% CI (-0.36; 0.06)] (Figure S4). The same result was observed between one and two years [SMD: -0.13, 95% CI (-0.48; 0.23)] and after three years [SMD: -0.24, 95% CI (-0.69; 0.22)] (Figure S4).

When only hormone-naïve TM studies were included, the absence of body fat gain remained up to one year of therapy compared to baseline [SMD: -0.20, 95% CI (-0.49; 0.08)], and a reduction was observed between one and two years [SMD: -0.41, 95% CI (-0.67; -0.15)] and after three years [SMD: -0.42, 95% CI (-0.78; -0.06)] (Figure S5).

Regarding puberty suppression, no temporal comparisons were possible. An increase in body fat mass was observed (SMD: 0.25, 95% CI [0.03; 0.46]) in TW who underwent puberty suppression, (Figure S6). Conversely, in TM who underwent puberty suppression, no differences in body fat mass were observed (SMD: 0.36, 95% CI [-0.05; 0.78]) (Figure S6).

Lean Mass

Of the 26 studies reporting this outcome for this comparison, 19 were included in the meta-analysis. Five studies were excluded because they either lacked extractable data (due to a graphical format) or did not report the duration of hormone therapy.

First, we conducted the analysis with all studies, regardless of whether the individuals were naïve or had undergone puberty suppression.

In TW, a decrease in lean mass was observed after one year of cross-sex hormone therapy compared to baseline [SMD: -0.22, 95% CI (-0.37; -0.08)] (Figure S7). No difference between two years and three years were observed when compared to baseline [SMD: -0.32, 95% CI (-0.76; 0.13)] (Figure S7).

When only studies with hormone-naïve TW were included, it was only possible to make a comparison after 1 year of therapy compared to baseline. A decrease in lean mass was observed [SMD: -0.26, 95% CI (-0.42; -0.09)] (Figure S8).

In TM, an increase in body lean mass was observed after one year of cross-sex hormone therapy compared to baseline [SMD: 0.51, 95% CI (0.39; 0.64)] (Figure S9). The same result was observed between one and two years [SMD: 0.39, 95% CI (0.24; 0.53)], and no differences were observed after three years when compared with baseline [SMD: 0.41, 95% CI (-0.09; 0.91)] (Figure S9).

When only hormone-naïve TM studies were included, an increase in body lean mass was observed after one year of hormone therapy compared to baseline [SMD: 0.55, 95% CI (0.39; 0.71)] (Figure S10). The same result was observed between one and two years [SMD: 0.44, 95% CI (0.27; 0.61)], and after three years [SMD: 0.57, 95% CI (0.21; 0.93)] (Figure S10).

Regarding puberty suppression, no temporal comparisons were possible. In both TW and TM who underwent puberty suppression, no significant differences were observed [SMD: 0.10, 95% CI (-0.44; 0.65) and SMD: -0.25, 95% CI (-0.62; 0.12)], from TW and TM, respectively) (Figure S11).

Upper-body strength

All 11 studies reporting this outcome for this comparison were included in the meta-analysis.

First, we conducted the analysis with all studies reporting this outcome, regardless of whether the individuals were naïve for hormone therapy. It was not possible to conduct analysis for puberty suppression.

In TW, a decrease in upper-body strength was observed after one year of cross-sex hormone therapy compared to baseline [SMD: -0.34, 95% CI (-0.54; -0.14)] (Figure S12). No difference between two years and three years were observed when compared to baseline [SMD: -0.22, 95% CI (-1.13; 0.70)] (Figure S12).

When only studies with hormone-naïve TW were included, a decrease in upper-body strength was observed after one year of cross-sex hormone therapy compared to baseline [SMD: -0.35, 95% CI (-0.56; -0.14)] (Figure S13). The same result was observed after follow-ups between two and three years [SMD: -0.67, 95% CI (-1.11; -0.22)] (Figure S13).

In TM, an increase in upper-body strength was observed after one year of cross-sex hormone therapy compared to baseline [SMD: 0.73, 95% CI (0.63; 0.83)] (Figure S14). A similar result was observed after follow-ups between one and two years [SMD: 0.91, 95% CI (0.35; 1.46)] and after three years when compared with baseline [SMD: 1.10, 95% CI (0.34; 1.87)] (Figure S14).

When only hormone-naïve TM studies were included, an increase in upper-body strength was observed after one year of cross-sex hormone therapy compared to baseline [SMD: 0.76, 95% CI (0.65; 0.86)] (Figure S15). Increases in strength were also observed after follow-ups between one and two years [SMD: 0.91, 95% CI (0.35; 1.46)], and no differences were observed after three years when compared with baseline [SMD: 1.03, 95% CI (0.00; 2.06)] (Figure S15).

Lower-body strength

The three studies reporting this outcome were included in the meta-analysis.

Analyses of lower-body strength were constrained by the limited number of available studies, precluding both temporal comparisons and assessments of puberty suppression - all studies included involved exclusively treatment-naïve patients. Cross-sex hormone therapy duration in TW and TM was 1 year in one study, ~2.5 years in the other and one did not reported duration. No significant differences were observed in TW when compared to baseline (i.e., before cross-sex hormone therapy) [SMD: -0.17, 95% CI [-0.58; 0.25]] (Figure S16). An increase in lower-body strength was observed in TM when compared to baseline [SMD: 0.48, 95% CI (0.28; 0.69)] (Figure S16).

VO₂ peak

Only one study reported VO₂ peak for these comparisons; therefore, a meta-analysis for this outcome was not performed.

Risk of bias

Quality scores ranged from 10 to 18, with most studies rated as high (66.7 %; N = 10) or fair (20.0 %; N = 3), and 13.3 % as low quality (N = 2) for cross-sectional studies. Using AXIS, the most common weaknesses identified in the articles were the lack of justification of the sample size (N = 13) and the use of a convenience sample (N = 14).

Among cohort and quasi-experimental studies, 65.6% exhibited moderate risk of bias, with limitations concerning confounding (15/32 studies, 46.8%, moderate risk) and selective reporting (22/32 studies, 68.8%, moderate risk). Bias in participant selection, intervention classification and deviations from intended interventions were predominantly low risk (24/32, 75%; 25/32, 78.1%; 30/32, 93.8%, respectively).

All RCTs were classified as some concerns. Methodological concerns primarily stemmed from selection of the reported result (3/3 studies, 100%, some concerns); deviation from intended interventions (3/3 studies, 100%, some concerns); and randomization process (2/3 studies, 66.6%, some concerns).

A detailed assessment of the risk of bias for individual studies is provided in the supplementary material (Table S3 and S4, Figure S17).

Discussion

Statement of principal findings

In addition to evaluating physical fitness and body composition differences between transgender and cisgender individuals, this systematic review and meta-analysis also aimed to examine the influence of GAHT duration on the outcomes, assess variations between TM and TW in response to therapy, compare naïve vs. non-naïve individuals, and explore the impact of puberty suppression on body composition and strength. Furthermore, the review evaluated the methodological rigor and risk of bias in existing studies to provide a comprehensive understanding of the evidence quality and practical implications for the findings.

The meta-analysis of 20 studies revealed that TW exhibited significantly higher fat mass than CM (SMD: 0.85), but similar to CW. In respect of lean mass, meta-analysis

of 21 studies revealed that TW showed higher values than CW (SMD: 0.98), but lower than CM (SMD: -1.17). Of relevance, the meta-analysis of 12 studies for upper-body strength and four for lower-body strength revealed that both were not significantly different between TW and CW (SMD: 0.41 and 0.05, respectively), but were markedly reduced in TW compared to CM (SMD: -1.79 and -1.99, respectively). Longitudinal hormone therapy (1–3 years) led to progressive increases in TW's fat mass (SMD: 0.44 to 0.70) and declines in both lean mass (SMD: -0.22 to -0.32) and upper-body strength (SMD: -0.34 to -0.67). VO₂ peak in TW did not differ from CW, but was lower than CM (SMD: -1.61).

The meta-analysis comparing TM and CM revealed significant differences, with TM showing intermediate body composition and strength metrics between CW and CM, even after an average therapy duration of 4 years.

Practical implications

This review shows that, despite TW exhibiting higher absolute lean mass compared to CW, no significant differences in upper- or lower-body strength were observed between the two groups after 1–3 years of hormone therapy. This finding challenges the assumption that potential residual lean mass inherently translates to functional strength advantages in this population. For example, TW's upper- (SMD: 0.41, 95% CI [-0.10; 0.92]) and lower-body strength (SMD: 0.05, 95% CI [-0.74; 0.83]) showed negligible divergence from CW, even as total body lean mass declined modestly in response to GAHT (SMD: -0.22 to -0.32). These results align with evidence [4] showing that, while TW retain higher absolute lean mass, body composition-adjusted strength metrics (e.g., relative strength per kg lean mass) converge with CW over time. This suggests that lean mass alone is an incomplete proxy for TW's athletic performance, as neuromuscular efficiency, training history, and fat distribution may play compensatory roles. In fact, when height-normalized, TW and CW appear to show comparable appendicular lean mass/height² or lean mass/height² [15, 66]. Importantly, a greater absolute lean mass not accompanied by increased functionality may actually impair performance, especially in weight-sensitive sports (e.g. cycling and climbing), which not surprisingly have the lightest athletes[67].

Furthermore, sport performance extends beyond physiology to encompass social, psychological, and cultural dimensions (e.g., stigmatization, discrimination, access to

sports opportunities, self-concept, self-esteem etc), which altogether may influence athletic engagement and achievement among transgender individuals. Indeed, evidence indicates that this population faces an elevated risk of adverse mental health outcomes [68], likely due to systemic stigma and discrimination across various contexts [69], including sports [70, 71]. While the extent to which potential “muscle memory” (i.e., long-lasting physiological effects of prior testosterone exposure) may counteract the influence of these psychosocial factors on athletic performance remains unestablished, the assumption of inherent competitive advantages for TW over CW does not appear to be robustly supported by existing evidence.

In fact, the absence of strength disparities between TW and CW found in the current review contradicts narratives framing male puberty as conferring irreversible athletic advantages despite GAHT. In a narrative review, Lundberg et al. (2024) [3] argue that male developmental traits (e.g., height, skeletal proportions) inherently disrupt fairness, yet the lack of measurable strength differences in the present systematic review suggests such claims may overemphasize structural factors while underestimating the impact of GAHT. For instance, TW’s VO₂ peak, when adjusted for weight, aligns with CW [4], further supporting parity in endurance capabilities. Furthermore, TW’s pre-therapy advantages in push-ups and sit-ups disappeared after 2 years of feminizing hormones among 46 individuals who started GAHT while in the United States Air Force [38]. These findings are corroborated by the current meta-analysis, endorsing nuanced, sport-specific policies rather than blanket bans.

Limitations of the available evidence and the review

This systematic review aligns with previous ones [1, 4] in highlighting critical research limitations. This includes the typically short study durations (<3 years) and a lack of data on elite athletes. Additionally, the potential conflation of trained and untrained individuals complicates extrapolation. The available evidence remains limited for specific outcomes (e.g., lower-body strength and VO₂ peak), particularly regarding RCTs examining the effects of GAHT on physical fitness and body composition (n = 3), as well as studies assessing the impact of puberty suppression (n = 6). Another literature weakness is the inconsistent reporting and adjustment for confounders, as few studies controlled for training history, diet, baseline fitness, physical activity and body composition or previous hormone therapy, potentially hindering the isolated effects of GAHT. Finally, there is very little literature involving transgender athletes of any age,

across all sport settings, and at any competitive level. Therefore, future studies must prioritize transgender athletes, assess sport-specific performance metrics, and evaluate long-term (e.g., >5 years) physiological and psychological changes, controlling for puberty suppression whenever possible.

The limitations of this review are related to the identified gaps in literature and include: the reliance on short-term assessments, limiting conclusions about the effects of GAHT on targeted outcomes in the long run; the heterogeneity of the studies assessed, variability in hormone regimens (e.g., types/doses of antiandrogens, estrogens), measurement methods (e.g., DEXA vs. MRI for body composition), and control groups (e.g., inconsistent physical activity tracking); the reliance on lean mass and strength as proxies for performance, rather than sport-specific outcomes (e.g., race times, power output), which limits practical relevance to real-world sport scenario; underrepresentation of puberty-suppressed cohorts, hampering the meta-analytic comparison between suppressed vs. non-suppressed individuals; and the inclusion of studies with risk of bias and with cross-sectional or retrospective designs.

Conclusions and perspectives

This systematic review and meta-analysis shows that, while TW exhibited higher absolute lean mass compared to CW, no significant differences in physical fitness metrics (i.e., upper-body strength, lower-body strength and VO₂ peak) were observed after 1–3 years of therapy. Although the current data do not justify blanket bans, critical gaps in literature were found, notably the underrepresentation of transgender athletes who may retain more “muscle memory”. Ideally, to resolve speculation, future long-term, longitudinal studies should prioritize performance-specific metrics in transgender athletes. However, one should be aware of the scarce number of transgender athletes, particularly in the elite sport, which complicates the feasibility of conducting powered studies involving high-performance transgender athletes within specific sport disciplines. In light of this context of imperfect evidence and despite the methodological challenges, continued research into physiological as well as psychosocial trajectories among transgender athletes with diverse demographics and clinical characteristics remains essential for developing equitable frameworks that balance justice, inclusion, and scientific rigor. Policies should remain dynamic, guided by evolving evidence and ethical imperatives, whereas acknowledging that fairness and non-discrimination are interdependent objectives necessitating nuanced, context-sensitive strategies.

References

1. Harper, J., et al., *How does hormone transition in transgender women change body composition, muscle strength and haemoglobin? Systematic review with a focus on the implications for sport participation*. Br J Sports Med, 2021. **55**(15): p. 865-872.
2. IOC. *IOC Framework Fairness Inclusion Non-Discrimination on Basis Gender Identity Sex Variations*. 2021 [cited 2023 August, 21]; Available from: <https://olympics.com/ioc/documents/athletes/ioc-framework-on-fairness-inclusion-and-nondiscrimination-on-the-basis-of-gender-identity-and-sex-variations>.
3. Lundberg, T.R., et al., *The International Olympic Committee framework on fairness, inclusion and nondiscrimination on the basis of gender identity and sex variations does not protect fairness for female athletes*. 2024. **34**(3): p. e14581.
4. Cheung, A.S., et al., *The Impact of Gender-Affirming Hormone Therapy on Physical Performance*. The Journal of Clinical Endocrinology & Metabolism, 2023. **109**(2): p. e455-e465.
5. Page, M.J., et al., *The PRISMA 2020 statement: an updated guideline for reporting systematic reviews*. 2021. **372**: p. n71.
6. Ouzzani, M., et al., *Rayyan—a web and mobile app for systematic reviews*. Systematic Reviews, 2016. **5**(1): p. 210.
7. Downes, M.J., et al., *Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS)*. 2016. **6**(12): p. e011458.
8. Sterne, J.A., et al., *ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions*. 2016. **355**: p. i4919.
9. Sterne, J.A.C., et al., *RoB 2: a revised tool for assessing risk of bias in randomised trials*. 2019. **366**: p. l4898.
10. Bagordo, A., et al., *Isokinetic Dynamometry as a Tool to Predict Shoulder Injury in an Overhead Athlete Population: A Systematic Review*. 2020. **8**(9): p. 124.
11. Yarali, M., et al., *RETRACTED: A systematic review of health care workers' knowledge and related factors towards burn first aid*. 2023. **20**(8): p. 3338-3348.
12. Yang, J., et al., *Cross-cultural differences in trait emotional intelligence: A meta-analysis*. Personality and Individual Differences, 2025. **241**: p. 113195.
13. Parvizi, A., et al., *RETRACTED: A systematic review of nurses' knowledge and related factors towards the prevention of medical device-related pressure ulcers*. 2023. **20**(7): p. 2843-2854.
14. Schwarzer, G., J.R. Carpenter, and G. Rücker *General Package for Meta-Analysis Version. Meta-Analysis with R*. 2020. 187-216.
15. Alvares, L.A., et al., *Body composition, exercise-related performance parameters and associated health factors of transgender women, cisgender women and cisgender men volleyball players*. 2025: p. bjsports-2024-108601.
16. Saitong, A., et al., *Physical Fitness and Exercise Performance of Transgender Women*. Medicine & Science in Sports & Exercise, 2025. **57**(1).
17. Sánchez Amador, L., et al., *Body composition and risk for sarcopenia in transgender women*. Nutrition, 2024. **123**: p. 112398.
18. Ceolin, C., et al., *Bone health and body composition in transgender adults before gender-affirming hormonal therapy: data from the COMET study*. J Endocrinol Invest, 2024. **47**(2): p. 401-410.

19. Ciancia, S., et al., *Effects of puberty suppression on bone, body composition, handgrip strength and glucolipid profile in early-pubertal transgender adolescents*. International Journal of Transgender Health: p. 1-13.
20. Hamilton, B., et al., *Strength, power and aerobic capacity of transgender athletes: a cross-sectional study*. Br J Sports Med, 2024. **58**(11): p. 586-597.
21. Pei, Q., et al., *Effects of gender-affirming hormone therapy on body fat: a retrospective case-control study in Chinese transwomen*. Lipids Health Dis, 2024. **23**(1): p. 146.
22. Boogers, L.S., et al., *Time Course of Body Composition Changes in Transgender Adolescents During Puberty Suppression and Sex Hormone Treatment*. J Clin Endocrinol Metab, 2024. **109**(8): p. e1593-e1601.
23. Chiccarelli, E., et al., *Fit Transitioning: When Can Transgender Airmen Fitness Test in Their Affirmed Gender?* Mil Med, 2023. **188**(7-8): p. e1588-e1595.
24. Tominaga, Y., et al., *Trans men can achieve adequate muscular development through low-dose testosterone therapy: A long-term study on body composition changes*. Andrology, 2025. **13**(2): p. 275-285.
25. Yaish, I., et al., *Sublingual Estradiol Offers No Apparent Advantage Over Combined Oral Estradiol and Cyproterone Acetate for Gender-Affirming Hormone Therapy of Treatment-Naive Trans Women: Results of a Prospective Pilot Study*. Transgend Health, 2023. **8**(6): p. 485-493.
26. Yamada, M., et al., *Effects of acute aerobic exercise on arterial stiffness in transgender men*. Front Physiol, 2023. **14**: p. 1294284.
27. Alvares, L.A.M., et al., *Cardiopulmonary capacity and muscle strength in transgender women on long-term gender-affirming hormone therapy: a cross-sectional study*. Br J Sports Med, 2022. **56**(22): p. 1292-1298.
28. Andrade, S.R.L., et al., *Bone mineral density, trabecular bone score and muscle strength in transgender men receiving testosterone therapy versus cisgender men*. Steroids, 2022. **178**: p. 108951.
29. Klaver, M., et al., *Change in Visceral Fat and Total Body Fat and the Effect on Cardiometabolic Risk Factors During Transgender Hormone Therapy*. J Clin Endocrinol Metab, 2022. **107**(1): p. e153-e164.
30. Bretherton, I., et al., *Insulin resistance in transgender individuals correlates with android fat mass*. Ther Adv Endocrinol Metab, 2021. **12**: p. 2042018820985681.
31. Gava, G., et al., *A Randomized Double-Blind Placebo-Controlled Pilot Trial on the Effects of Testosterone Undecanoate Plus Dutasteride or Placebo on Muscle Strength, Body Composition, and Metabolic Profile in Transmen*. J Sex Med, 2021. **18**(3): p. 646-655.
32. Navabi, B., et al., *Pubertal Suppression, Bone Mass, and Body Composition in Youth With Gender Dysphoria*. Pediatrics, 2021. **148**(4).
33. Nokoff, N.J., et al., *Body Composition and Markers of Cardiometabolic Health in Transgender Youth on Gonadotropin-Releasing Hormone Agonists*. Transgend Health, 2021. **6**(2): p. 111-119.
34. Yun, Y., D. Kim, and E.S. Lee, *Effect of Cross-Sex Hormones on Body Composition, Bone Mineral Density, and Muscle Strength in Trans Women*. J Bone Metab, 2021. **28**(1): p. 59-66.
35. Gava, G., et al., *A comparison of 5-year administration of cyproterone acetate or leuprolide acetate in combination with estradiol in transwomen*. Eur J Endocrinol, 2020. **183**(6): p. 561-569.

36. Jenkins, C.L., et al., *Performance in Transgender Females Versus Cisgender Males and Females*. 2020. **23**(6).
37. Nokoff, N.J., et al., *Body Composition and Markers of Cardiometabolic Health in Transgender Youth Compared With Cisgender Youth*. *J Clin Endocrinol Metab*, 2020. **105**(3): p. e704-14.
38. Roberts, T.A., J. Smalley, and D. Ahrendt, *Effect of gender affirming hormones on athletic performance in transwomen and transmen: implications for sporting organisations and legislators*. *British Journal of Sports Medicine*, 2020: p. bjsports-2020-102329.
39. Shadid, S., et al., *Effects of Gender-Affirming Hormone Therapy on Insulin Sensitivity and Incretin Responses in Transgender People*. *Diabetes Care*, 2020. **43**(2): p. 411-417.
40. van Velzen, D.M., et al., *Variation in sensitivity and rate of change in body composition: steps toward individualizing transgender care*. *Eur J Endocrinol*, 2020. **183**(5): p. 529-536.
41. Chrisostomo, K.R., et al., *Transwomen and bone mineral density: a cross-sectional study in Brazilian population*. *Br J Radiol*, 2020. **93**(1111): p. 20190935.
42. Scharff, M., et al., *Change in grip strength in trans people and its association with lean body mass and bone density*. *Endocr Connect*, 2019. **8**(7): p. 1020-1028.
43. Wiik, A., et al., *Muscle Strength, Size, and Composition Following 12 Months of Gender-affirming Treatment in Transgender Individuals*. *J Clin Endocrinol Metab*, 2020. **105**(3).
44. Auer, M.K., et al., *Effects of Sex Hormone Treatment on the Metabolic Syndrome in Transgender Individuals: Focus on Metabolic Cytokines*. *J Clin Endocrinol Metab*, 2018. **103**(2): p. 790-802.
45. Figuera, T.M., et al., *Impact of cross-sex hormone therapy on bone mineral density and body composition in transwomen*. *Clin Endocrinol (Oxf)*, 2018. **88**(6): p. 856-862.
46. Gava, G., et al., *Testosterone undecanoate and testosterone enanthate injections are both effective and safe in transmen over 5 years of administration*. *Clin Endocrinol (Oxf)*, 2018. **89**(6): p. 878-886.
47. Klaver, M., et al., *Early Hormonal Treatment Affects Body Composition and Body Shape in Young Transgender Adolescents*. *J Sex Med*, 2018. **15**(2): p. 251-260.
48. Tack, L.J.W., et al., *Proandrogenic and Antiandrogenic Progestins in Transgender Youth: Differential Effects on Body Composition and Bone Metabolism*. *J Clin Endocrinol Metab*, 2018. **103**(6): p. 2147-2156.
49. Hannema, S.E., et al., *Efficacy and Safety of Pubertal Induction Using 17 β -Estradiol in Transgirls*. *J Clin Endocrinol Metab*, 2017. **102**(7): p. 2356-2363.
50. Klaver, M., et al., *Cross-sex hormone therapy in transgender persons affects total body weight, body fat and lean body mass: a meta-analysis*. *Andrologia*, 2017. **49**(5).
51. Auer, M.K., et al., *12-months metabolic changes among gender dysphoric individuals under cross-sex hormone treatment: a targeted metabolomics study*. *Sci Rep*, 2016. **6**: p. 37005.
52. Van Caenegem, E., et al., *Preservation of volumetric bone density and geometry in trans women during cross-sex hormonal therapy: a prospective observational study*. *Osteoporos Int*, 2015. **26**(1): p. 35-47.
53. Van Caenegem, E., et al., *Body composition, bone turnover, and bone mass in trans men during testosterone treatment: 1-year follow-up data from a prospective case-controlled study (ENIGI)*. *Eur J Endocrinol*, 2015. **172**(2): p. 163-71.

54. Klink, D., et al., *Bone mass in young adulthood following gonadotropin-releasing hormone analog treatment and cross-sex hormone treatment in adolescents with gender dysphoria*. J Clin Endocrinol Metab, 2015. **100**(2): p. E270-5.
55. Pelusi, C., et al., *Effects of three different testosterone formulations in female-to-male transsexual persons*. J Sex Med, 2014. **11**(12): p. 3002-11.
56. Vilas, M.V.A., et al., *Nutritional Status and Obesity Prevalence in People with Gender Dysphoria*. AIMS Public Health, 2014. **1**(3): p. 137-146.
57. Wierckx, K., et al., *Cross-sex hormone therapy in trans persons is safe and effective at short-time follow-up: results from the European network for the investigation of gender incongruence*. J Sex Med, 2014. **11**(8): p. 1999-2011.
58. Van Caenegem, E., et al., *Low bone mass is prevalent in male-to-female transsexual persons before the start of cross-sex hormonal therapy and gonadectomy*. Bone, 2013. **54**(1): p. 92-7.
59. Van Caenegem, E., et al., *Bone mass, bone geometry, and body composition in female-to-male transsexual persons after long-term cross-sex hormonal therapy*. J Clin Endocrinol Metab, 2012. **97**(7): p. 2503-11.
60. Mueller, A., et al., *Effects of intramuscular testosterone undecanoate on body composition and bone mineral density in female-to-male transsexuals*. J Sex Med, 2010. **7**(9): p. 3190-8.
61. Mueller, A., et al., *Body composition and bone mineral density in male-to-female transsexuals during cross-sex hormone therapy using gonadotrophin-releasing hormone agonist*. Exp Clin Endocrinol Diabetes, 2011. **119**(2): p. 95-100.
62. Lapauw, B., et al., *Body composition, volumetric and areal bone parameters in male-to-female transsexual persons*. Bone, 2008. **43**(6): p. 1016-21.
63. Meriggiola, M.C., et al., *Effects of testosterone undecanoate administered alone or in combination with letrozole or dutasteride in female to male transsexuals*. J Sex Med, 2008. **5**(10): p. 2442-53.
64. Elbers, J.M., et al., *Effects of sex steroid hormones on regional fat depots as assessed by magnetic resonance imaging in transsexuals*. Am J Physiol, 1999. **276**(2): p. E317-25.
65. Elbers, J.M., et al., *Long-term testosterone administration increases visceral fat in female to male transsexuals*. J Clin Endocrinol Metab, 1997. **82**(7): p. 2044-7.
66. Alvares, L.A.M., et al., *Cardiopulmonary capacity and muscle strength in transgender women on long-term gender-affirming hormone therapy: a cross-sectional study*. 2022. **56**(22): p. 1292-1299.
67. Wagner, D.R., et al., *Multicomponent body composition of university club sport athletes*. Journal of the International Society of Sports Nutrition, 2025. **22**(1): p. 2446575.
68. Tagg, B., *Transgender Netballers: Ethical Issues and Lived Realities* %J Sociology of Sport Journal. 2012. **29**(2): p. 151-167.
69. Hatzenbuehler, M.L. and J.E. Pachankis, *Stigma and Minority Stress as Social Determinants of Health Among Lesbian, Gay, Bisexual, and Transgender Youth: Research Evidence and Clinical Implications*. Pediatric Clinics of North America, 2016. **63**(6): p. 985-997.
70. López-Cañada, E., et al., *Physical Activity and Sport in Trans Persons Before and After Gender Disclosure: Prevalence, Frequency, and Type of Activities* %J Journal of Physical Activity and Health. 2020. **17**(6): p. 650-656.
71. Kroshus, E., et al., *Improving inclusion and well-being of trans and gender nonconforming collegiate student-athletes: foundational concepts from the National*

Collegiate Athletic Association Summit on Gender Identity and Student–Athlete Participation. 2023. 57(10): p. 564-570.

Table 1 - Methodological characteristics of the studies included

Author (date)	Population				Age* (mean±SD or Range)				Comparison	Outcomes	Design	Physical activity experience/practice	
	N	Transgender		Cisgender		Transgender		Cisgender					
		TW	TM	CW	CM	TW	TM	CW					CM
Alvares et al (2025)	23	7	-	8	8	30 (28-33)	-	26 (22-29)	28 (25-31)	TW vs. CW TW vs. CM	FM/LM Handgrip/CMJ Vo ₂ max	CS	Amateur volleyball players Hours/week TW: 4(3-5) CW: 14(11-16) CM: 4(3-6)
Saitong et al (2025)	60	30	-	15	15	GAS 27±4 without GAS 30±4	-	29±5	28±5	TW vs. CW TW vs. CM	FM/LM Elbow extension (peak torque) /CMJ Vo ₂ max	CS	MET min/week TW with GAS 2169±1062 TW without GAS: 2343 ± 1056 CW: 2193 ± 770 CM: 2563 ± 784
Amador et al (2024)	105	37	-	34	34	29 (20-38)	-	29 (20-42)	29 (21-36)	TW vs. CW TW vs. CM	FM/LM	CS	NR
Ceolin et al (2024)	101	26	29	26	20	26±7	24±6	24±4	23±6	TW vs. CW TM vs. CM	FM/LM Handgrip	CS	MET min/week TW: ~1755 TM: ~4377 CW: ~2505 CM: ~1770
Ciancia et al (2024)	46	26	20	-	-	NR	NR	-	-	Pre vs. post	FM/LM	Retrospective	NR
Hamilton et al (2024)	75	23	12	21	19	34±10	34±7	30±9	37±9	TW vs. CW TW vs. CM TM vs. CW TM vs. CM	FM/LM Handgrip/ACJ	CS	Play a Competitive Sport or Physically train at least 3x per week
Pei et al (2024)	99	GAHT59 not GAHT 40	-	-	-	on GAHT 23±3 not on GAHT 25±5	-	-	-	Pre vs. post	FM/LM	Retrospective	NR
Boogers et al (2023)	463	154	309	-	-	14±2	14±2	-	-	Pre vs. post	FM/LM	Retrospective	NR
Chiccarelli et al (2023)	374	228	146	146	228	27	26	26	27	TW vs. CW TW vs. CM	Push-ups 1.5-mile run	Retrospective	NR

Tominaga et al (2023)	291	-	HD 103 LD 188	-	-	-	25±6	-	-	Pre vs. post	LM/FM Handgrip	Retrospective	NR
Yaish et al (2023)	22	CO 11 SLE 11	-	-	-	SLE 26±6 CO 20±2	-	-	-	Pre vs. post	FM/LM	Prospective	NR
Yamada et al (2023)	36	-	12	12	12	-	27±4	25±3	25±2	TM vs. CW TM vs. CM	FM/LM	CS	NR
Alvares et al (2022)	42	15	-	13	14	34±5	-	36±4	37±4	TW vs. CW TW vs. CM	FM/LM Handgrip	CS	TW 3 insufficiently active A 2 insufficiently active B 8 active 2 very active CW 2 insufficiently active A 4 active 7 very active CM 2 sedentary 1 insufficiently active A 7 active 4 very active* ~5 x week moderate/vigorous in both groups
Andrade et al (2022)	38	-	19	-	19	-	24±3	-	24±3	TM vs. CM	FM/LM Handgrip	CS	NR
Klaver et al (2022)	341	179	162	-	-	29 (23-43)	24 (21-33)	-	-	Pre vs. post	FM/LM	Prospective	NR
Bretherton et al (2021)	162	41	43	48	30	41 (26-53)	29 (25-33)	28 (24-39)	32 (26-41)	TW vs. CW TW vs. CM TM vs. CW TM vs. CM	FM/LM	CS	NR
Gava et al (2021)	14	-	TU+PL 7 TU+DT 7	-	-	-	TU+PL 35±6 TU+DT 36±6	-	-	Pre vs. post	FM/LM	RCT	NR
Navabi et al (2021)	170	51	119	-	-	15±2	15±2	-	-	Pre vs. post	FM/LM	Retrospective	NR
Nokoff et al (2021)	48	8	9	14	17	14±1	14±2	14±2	14±1	TW vs. CW TW vs. CM	FM/LM	CS	NR

										TM vs. CW TM vs. CM			
Yun et al (2021)	11	11	-	-	-	29 (20-47)	-	-	-	Pre vs. post	FM/LM Handgrip	Prospective	NR
Gava et al (2020)	50	CPA+E 25 Leu+E 25	-	-	-	CPA+E: 31±9 Leu+E: 33±13	-	-	-	Pre vs. post	FM	Prospective	NR
Jenkins et al (2020)	54	18	-	18	18	31±6	-	NR	NR	TW vs. CW TW vs. CM	FM/LM VJH/Push-up	CS	MET min/week 1930-5421
Nokoff et al (2020)	143	14	21	23/ 42	24/ 19	16±1	17±1	16±1/ 15±2	16±1/ 15±2	TW vs. CW TW vs. CM TM vs. CW TM vs. CM	FM/LM	CS	NR
Roberts et al (2020)	75	46	29	NR	NR	27±5	26±7	-	-	TW vs. CW TW vs. CM TM vs. CW TM vs. CM	Push-ups/Sit-ups 1.5-mile run	Retrospective	Regular exercise. Type and intensity vary by occupation.
Shadid et al (2020)	90	55	35	-	-	34±2	26±1	-	-	Pre vs. post	FM/LM	Prospective	NR
Velzen et al (2020)	611	288	323	-	-	25 (21-38)	21 (19-25)	-	-	Pre vs. post	FM/LM	Prospective	NR
Chrisostomo et al (2019)	62	31	-	31	31	29 (27-31)	-	30 (27-37)	29 (26-32)	TW vs. CM vs. CW	FM/LM	CS	TW: 54.8% active CW: 9.6% active CM: 0% active
Scharff et al (2019)	527	249	278	-	-	28 (23-40)	23 (20-30)	-	-	Pre vs. post	LM Handgrip	Prospective	NR
Wiik et al (2019)	23	11	12	-	-	27±4	25±5	-	-	Pre vs. Post	Isometric dynamometer	Prospective	TW 236 ± 206 baseline 230±163 12 months TM min/week 216±67 baseline 246±128 12 months
Auer et al (2018)	69	45	24	-	-	35±1	28±1	-	-	Pre vs. post	FM/LM	Prospective	NR
Fighera et al (2018)	181	142		17	22	34±10	-	33±5	31±6	Pre vs. post TW vs. CW TW vs. CM	FM	Prospective	NR
Gava et al (2018)	50	-	TU 25 TE	-	-	-	TU 30±5 TE	-	-	Pre vs. post	FM/LM	Retrospective	NR

25						30±7							
Klaver et al (2018)	192	71	121	-	-	start of GnRHa 15±2 start of CHT 16±1	-	-	-	Pre vs. post	FM/LM	Retrospective	NR
Tack et al (2018)	65	21	44	-	-	16±1	16±1	-	-	Pre vs. post	FM/LM Handgrip	Prospective	NR
Hannema et al (2017)	28	28	-	-	-	16 (14-19)	-	-	-	Pre vs. post	FM/LM	Prospective	NR
Klaver et al (2017)	341	179	162	-	-	29 (18-66)	24 (18-58)	-	-	Pre vs. post	FM	Prospective	NR
Auer et al (2016)	40	20	20	-	-	NR	NR	-	-	Pre vs. post	FM/LM Handgrip	Prospective	NR
Caenegem et al (2015a)	49	49	-	-	49	33±12	-	-	-	Pre vs. post	FM/LM Handgrip	Prospective	Sport index TW: 3 ± 1 CM: 3 ± 1
Caenegem et al (2015b)	23	-	23	23	-	-	27±9	27±9	-	TM vs. CW Pre vs. post	FM/LM Handgrip	Prospective	Baecke's questionnaire TM: 9±2 CW 9±2
Klink et al (2015)	34	15	19	-	-	start of GnRHa 15±2 start of CH 17±1	start of GnRHa 15±2 start of CHT 16±2	-	-	Pre vs. post	FM/LM	Prospective	NR
Pelusi et al (2014)	45	-	45	-	-	-	TD 31 (28-34) T-gel 29 (27-32) TU 28 (26-31)	-	-	Pre vs. post	FM/LM	RCT	NR
Vilas et al (2014)	157	91	66	-	-	35	30	-	-	Pre vs. post	FM	Quasi-experimental	Sedentary PAL < 2 (1-2)
Wierckx et al (2014)	106	Ghent 47 Oslo	Ghent 27 Oslo	-	-	Ghent Unit 32±15	Ghent Unit 27±9	-	-	Pre vs. post	FM/LM	Prospective	NR

	6	26				Oslo 19±2	Oslo Unit 22±5						
Caenegem et al (2013)	50	25	-	-	23	37 (28–42)	-	36 (28–42)	-	TW vs. CM	FM/LM Handgrip	CS	Baecke's questionnaire TW: 8(7–9) CM: 8(8–9)
Caenegem et al (2012)	100	-	50	50	-	-	37±8	-	38±8	TM vs. CW	FM/LM Handgrip	CS	NR
Mueller et al (2010a)	45	-	45	-	-	-	30±9	-	-	Pre vs. post	FM/LM	Prospective	NR
Mueller et al (2010b)	84	84	-	-	-	36±11	-	-	-	Pre vs. post	FM/LM	Prospective	NR
Lapauw et al (2008)	69	23	-	-	46	41±7	-	-	40±7	TW vs. CM	FM/LM Handgrip	CS	Baecke's questionnaire TW: 2±1 CM:3±1
							TU 34±4						
Meriggiola et al (2008)	15	-	15	-	-	-	TU+D 36±5	-	-	Pre vs. post	FM/LM	RCT	NR
							TU+L 37±6						
Elbers et al (1999)	40	20	20	-	-	26±7	24±7	-	-	Pre vs. post	FM/LM	Prospective	NR
Elbers et al (1997)	10	-	10	10	-	-	24 (16-33)	-	24 (16-33)	Pre vs. post TM vs. CW	FM/LM	Prospective	NR

Legend: CS: cross-sectional; TW: trans women; TM: trans men; CW: cis women; CM: cis men; FM: fat mass; LM: lean mass; NR: not reported; GnRHα: gonadotropin-releasing hormone agonists; CHT: cross-sex hormonal treatment; GAHT: gender-affirming hormone therapy; GAS: gender-affirming surgery; HT: hormone therapy; LD: low dose of testosterone; HD: high dose of testosterone. *Physical activity Level measured by the short-IPAQ version; TE: testosterone enanthate; TU: testosterone undecanoate; CPA: cyproterone acetate; E: estradiol; Leu: leuprolide acetate; DT: dutasteride; PL: placebo; ACJ: Absolute counter movement jump; VJH: vertical jump height; SLE: estradiol only protocol; CO: combined oral; TD: testosterone depot; T-gel: testosterone gel; RCT: randomized clinical trial

Figures

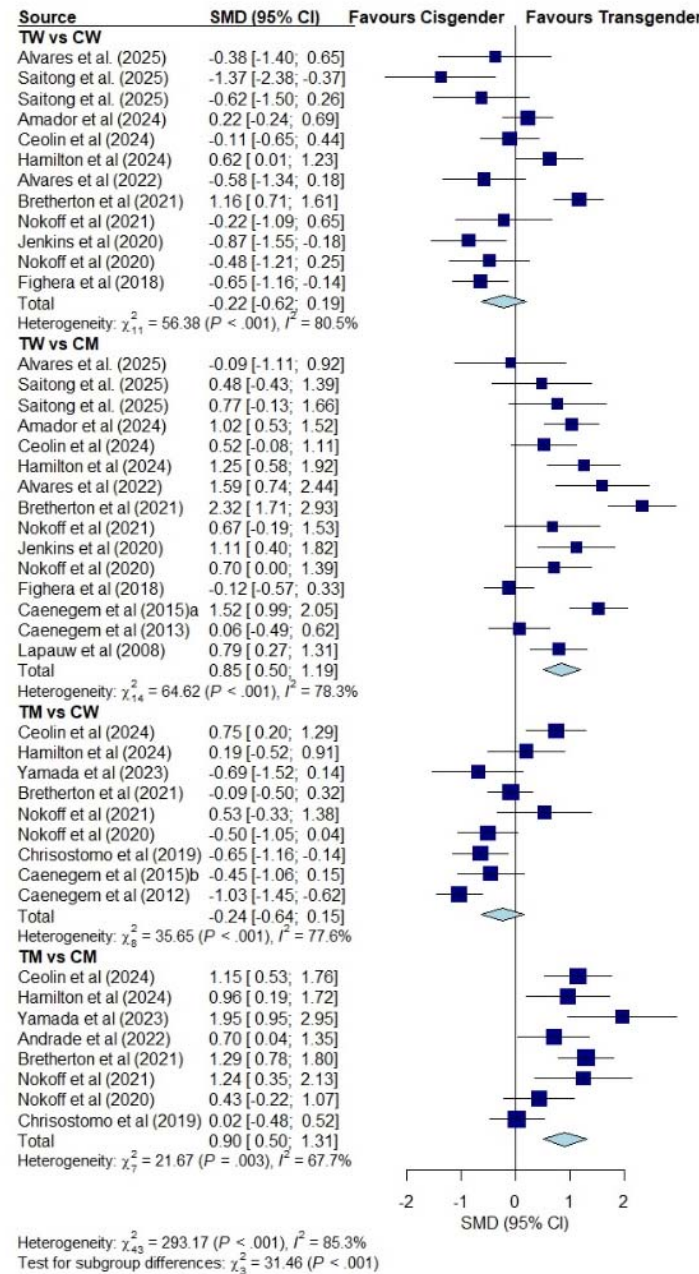


Figure 1. Forest plot: Fat mass in transgender women vs. cisgender women, transgender women vs. cisgender men, transgender men vs. cisgender women, and transgender men vs. cisgender men.

Legend: TW: transgender women; TM: transgender men; CW: cisgender women; CM: cisgender men; SMD: standard mean difference; CI: confidence interval.

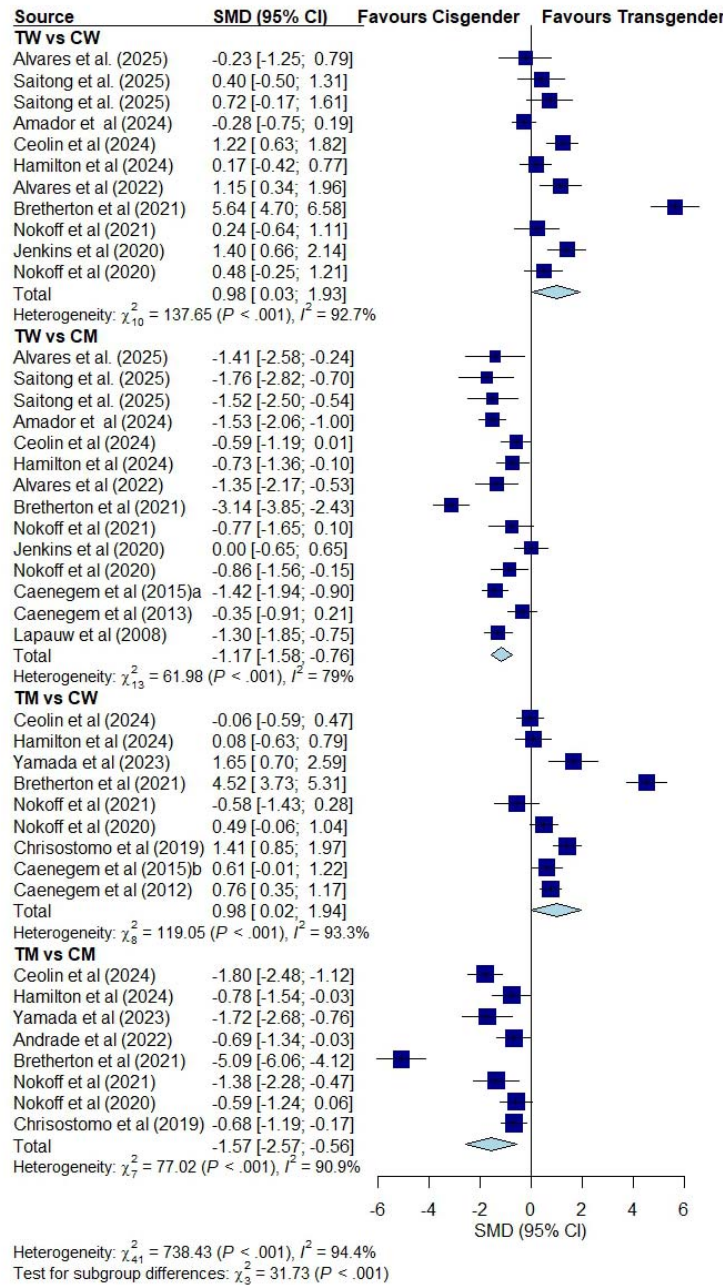


Figure 2. Forest plot: Lean mass in transgender women vs. cisgender women, transgender women vs. cisgender men, transgender men vs. cisgender women, and transgender men vs. cisgender men.

Legend: TW: transgender women; TM: transgender men; CW: cisgender women; CM: cisgender men; SMD: standard mean difference; CI: confidence interval.

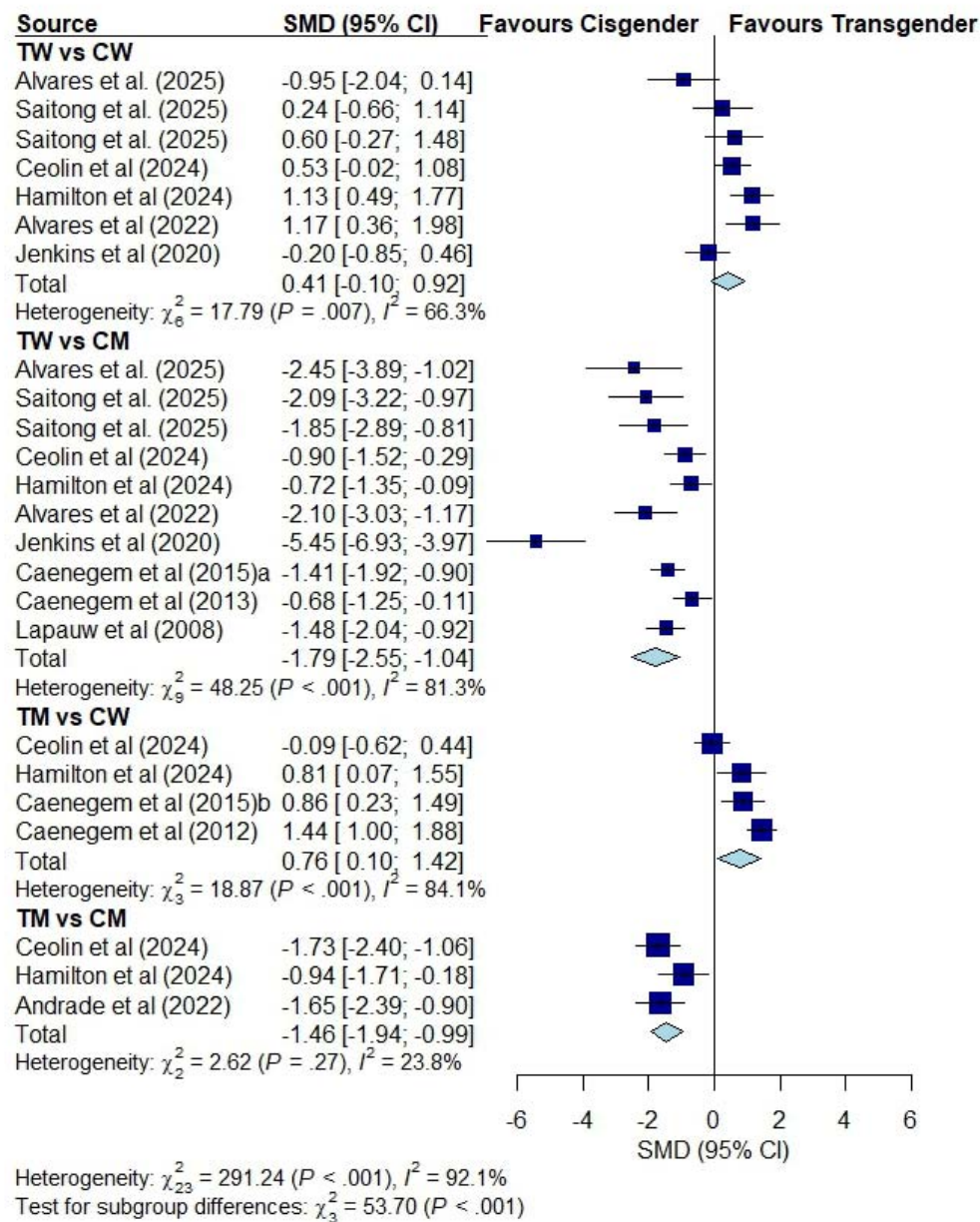


Figure 3. Forest plot: Upper-body strength in transgender women vs. cisgender women, transgender women vs. cisgender men, transgender men vs. cisgender women, and transgender men vs. cisgender men.

Legend: TW: transgender women; TM: transgender men; CW: cisgender women; CM: cisgender men; SMD: standard mean difference; CI: confidence interval.

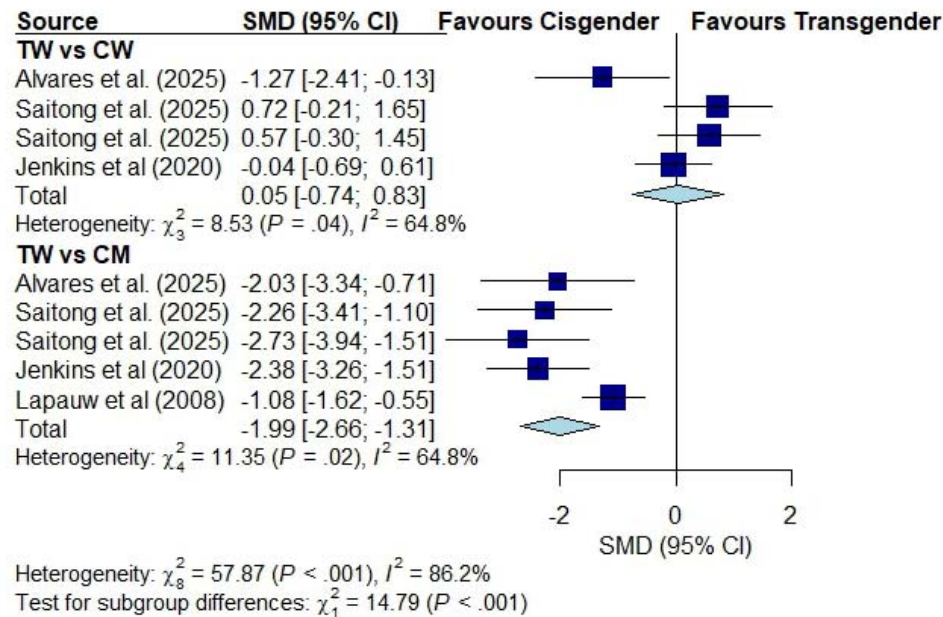


Figure 4. Forest plot: Lower-body strength in transgender women vs. cisgender women and transgender women vs. cisgender men.

Legend: TW: transgender women; TM: transgender men; CW: cisgender women; CM: cisgender men; SMD: standard mean difference; CI: confidence interval.

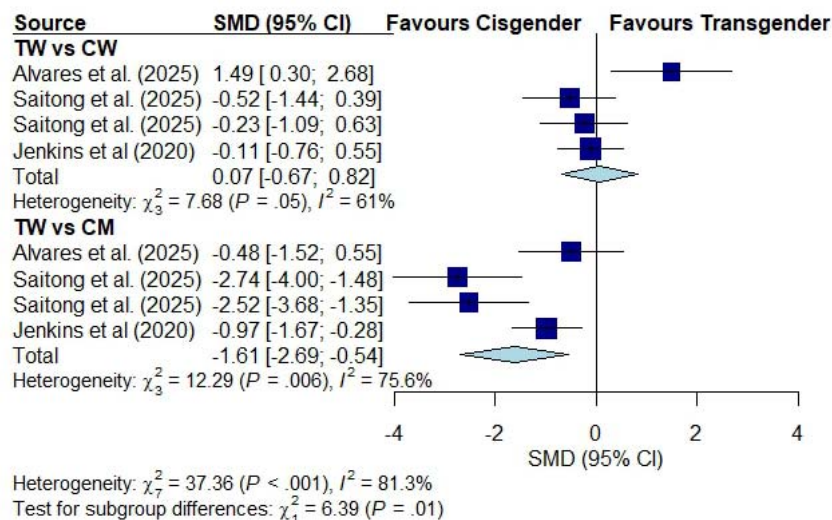


Figure 5. Forest plot: VO₂ peak in transgender women vs. cisgender women and transgender women vs. cisgender men.

Legend: TW: transgender women; TM: transgender men; CW: cisgender women; CM: cisgender men; SMD: standard mean difference; CI: confidence interval.